LINE-1 retrotransposons in cancer: Investigating human cervical carcinogenesis and developing an innovative transgenic mouse model for measuring somatic LINE-1 activity

Summary of the Ph.D. Thesis

Réka Karkas, MD

Supervisor: Lajos Mátés, Ph.D.

Doctoral School of Multidisciplinary Medical Sciences

Albert Szent-Györgyi Medical School

University of Szeged

Szeged

2025

List of publications

The publications served as the basis of the Ph.D. thesis:

I. **Réka Karkas**, Khaloon Sadiq Ahmed Abdullah*, László Kaizer, Ádám Ürmös, Lilla Tiszlavicz, Tibor Pankotai, Istvan Nagy, Lajos Mátés, Farkas Sükösd:
LINE-1 ORF1p is a Promising Biomarker in Cervical Intraepithelial Neoplasia Degree Assessment **International. Journal of Gynecological Pathology (2024)** DOI: 10.1097/PGP.0000000000001035 (Q2)

II. Anna Georgina Kopasz, Dávid Zsolt Pusztai, **Réka Karkas**, Liza Hudoba, Khaldoon Sadiq Ahmed Abdullah, Gergely Imre, Gabriella Pankotai-Bodó, Ede Migh, Andrea Nagy, András Kriston, Péter Germán, Andrea Bakné Drubi, Anna Molnár, Ildikó Fekete, Virág Éva Dani, Imre Ocsovszki, László Géza Puskás, Péter Horváth, Farkas Sükösd, Lajos Mátés
A versatile transposon-based technology to generate loss- and gain-of-function phenotypes in the mouse liver **BMC BIOLOGY (2022)** 20: 1 Paper: 74, 17 p. DOI: 10.1186/s12915-022-01262-x (D1)

III. Lajos Mátés, **Réka Karkas**, Sadiq Ahmed Abdullah Khaldoon, Andrea Mátés-Nagy, Gergely Imre, Anna Georgina Kopasz, Ildikó Fekete, Péter Horváth, Farkas Sükösd
Measurement of somatic Ll retrotransposition activity
Patent (2023) International Application Number: PCT/HU2023/050014

Other publications published during the Ph.D. scholarship:

IV. Gergely Imre, Bertalan Takács, Erik Czipa, Andrea Bakné Drubi, Gábor Jaksa, Dóra Latinovics, Andrea Nagy, **Réka Karkas**, Liza Hudoba, Bálint Márk Vásárhelyi, Gabriella Pankotai-Bodó, András Blastyák, Zoltán Hegedűs, Péter Germán, Balázs Bálint, Khaldoon Sadiq Ahmed Abdullah, Anna Georgina Kopasz, Anita Kovács, László G. Nagy, Farkas Sükösd, Lajos Mátés *Prolonged activity of the transposase helper may raise safety concerns during DNA transposon-based gene therapy.* **Molecular Therapy Methods Clin. Dev. (2023)** 29 pp. 145-159. , 15 p. DOI: 10.1016/j.omtm.2023.03.003 *(Q1)*

V. Lajos Mátés, Anna Georgina Kopasz, **Réka Karkas**, Khaloon Sadiq Ahmed Abdullah, Andrea Mátés-Nagy, Gergely Imre, Ildikó Fekete, Péter Horváth, Farkas Sükösd:
Versatile vector system for directed gene expression modifications
Patent (2023) International Application Number: PCT/HU2023/050013

1 INTRODUCTION

Once considered "junk" DNA, transposable elements (TEs) were historically underappreciated. They were presumed to be unimportant, and their high copy number and repetitive nature caused unique technical challenges upon investigation since they occupy almost half (~45%) of the human genome. However, recent advancements in genomic technologies and growing interest in the human mobilome have revealed significant roles in genome evolution, diversity, and disease. These sequences are now recognized for their roles in shaping the genome and contributing to various pathologies, including cancer, highlighting their importance in human health and disease.

LINE-1 retrotransposons are the only active transposable elements in the human genome. These elements encode ORF1 and ORF2 proteins, which are essential for their “copy-and-paste” self-propagation. These elements entered the genomes of mammalian ancestors around 70 million years ago and have since successfully co-evolved with us. LINE-1 elements now constitute 17% of the human genome, with around 500,000 total copies, of which 100-150 are still retrotransposition-competent. LINE-1 retrotransposons have a role in genomic evolution by fostering genetic diversity through mechanisms like creating gene duplications, exon shuffling, and alternative splicing by inserting themselves into different genomic locations. LINE-1 insertions can disrupt gene function, alter gene expression, and contribute to large-scale genomic rearrangements. Germline retrotransposition events can lead to heritable genomic variations and contribute to evolution. Dysregulated somatic retrotransposition can result in genomic instability, potentially leading to cancer.

To maintain genomic integrity, multiple cellular defense mechanisms regulate LINE-1 activity at various stages of its life cycle. If these defense mechanisms are compromised, LINE-1 activation can initiate tumor evolution. Environmental exposures, such as therapeutic drugs, radiation, and chemical pollutants, might impair these mechanisms, leading to LINE-1 activation and subsequent tumorigenesis. Epidemiological studies suggest that environmental factors play a more significant role in sporadic cancers than genetic factors. To date, investigating this subject has posed considerable difficulties. Quantifying LINE-1 retrotransposition events presents significant challenges due to the abundance and repetitive nature of these elements. Various assays have been developed, however, studying LINE-1 retrotransposition in physiological contexts remains difficult.

The ORFeus reporter system was developed as a specific tool for quantifying LINE-1

retrotransposition. It utilizes a synthetic LINE-1 element linked to an EGFP reporter gene inserted into the 3′ UTR of the optimized LINE-1 sequence, but in an antisense direction to the LINE-1 itself. An intronic sequence is inserted in sense direction to interrupt the reporter gene. In order to remove the intron, the LINE-1 RNA needs to undergo splicing. After that, if the reconstituted reporter is reverse transcribed, integrated into the genome, and antisense expression occurs, the reassembled reporter can be visualized. The retrotransposed reporter exhibits strong and permanent EGFP expression in the affected cell and in all its daughter cells, providing a clear and lasting marker for retrotransposition events.

The use of tumor-derived cell lines to study somatic L1 activity, while accessible, may not fully represent *in vivo* conditions, as the defense mechanisms that normally suppress L1 activity are often only partially functional in long-cultured tumor cell lines. Another strategy is the use of conventional germline transgenic mice. However, this approach is limited by early embryonic retrotransposition, which leads to ubiquitous expression of the reporter gene in all progeny cells, making it hard to distinguish between germline and somatic retrotransposition.

To study somatic L1 retrotransposition in a more physiological setting, we propose the development of an ORFeus reporter-bearing, somatically transgenic mouse model by harnessing our DNA transposon-based, therapeutic gene delivery-coupled, somatic transgenesis method.

2 AIMS OF THE THESIS

Despite considerable progress in cancer research, understanding the exact causes and mechanisms behind sporadic cancer development remains an enormous challenge. This thesis focuses on a lesser-explored area of cancer biology: the involvement of LINE-1 retrotransposons in the initiation of cancer. LINE-1 sequences are the most abundant group of mobile genetic elements in the human genome, and their uncontrolled activity acts as a somatic mutator.

In the present thesis, our aims were to:

i) Investigate the expression of the LINE-1 ORF1 protein in human cervical specimens during the different stages of cervical carcinogenesis and evaluate its potential as a biomarker for the diagnosis of cervical intraepithelial neoplasia, by setting up a new scoring system to

evaluate LINE-1 ORF1p immunostainings of human uterine cervical epithelia, incorporating both the staining intensity and the extent of epithelial involvement,

ii) to develop a novel technological platform to overcome the problems of the existing assays investigating somatic retrotransposition rate by establishing the first somatically transgenic mouse model for monitoring LINE-1 activity,

iii) furthermore, to utilize this technological platform to investigate unconventional genotoxic effects of chemicals on somatically transgenic hepatocytes of the LINE-1 reporter-bearing mouse and to validate the method as a chemical risk assessment technology, enhancing the currently available toxicological toolkit.

3 Materials and Methods

3.1. Preparation of the Human Tissue Specimens: Archived pathological samples were used to create tissue microarrays (TMAs) encompassing 143 human FFPE specimens: CIN I (n = 20), CIN II (n = 46), CIN III (n = 14), and cervical cancer (n = 32), controls included 31 non-dysplastic cervical tissues (normal n = 24, atrophic n = 7). Tumor specimens were assessed for grade, size, lymph node involvement, and metastasis status. Patient ages ranged from 22 to 83 years (mean age 52), with samples collected from 2018 to 2022 at the University of Szeged, Hungary. The study was approved by the Institutional Committee of Science and Research Ethics (Ref: BM/3049/2023).

3.2. Animal Care and Maintenance: Mice were bred in the Central Animal House at the Biological Research Centre, Szeged, Hungary, with pathogen-free status confirmed quarterly. They were maintained under a 12-hour light/dark cycle at 22 °C with *ad libitum* access to food and water. All procedures followed protocols approved by the Institutional Animal Care and Use Committee (Ethical authorization number: XVI./801/2018). The Fah mutant strain C57BL/6N$^{-Fahtm1(NCOM)Mfgc/Biat}$ was used, with Fah$^{-/-}$ mice receiving Orfadin® in drinking water.

3.3. Immunohistochemistry: Immunostaining was performed on the Bond Max Autostainer using the Bond Polymer Detection System. Antibodies used for human cervix FFPE specimens included anti-LINE-1 ORF1p, anti-p16, and anti-Ki67.

Mice were sacrificed at 3 months post-injection. Liver tissues were collected, overnight fixed in 4% formalin, embedded in paraffin, and sectioned into 5-μm slices. The EnVision FLEX Mini Kit (DAKO) was used to perform the IHC protocol with antigen retrieval conducted in a PT Link machine (DAKO). Primary antibodies included: rabbit polyclonal anti-FAH, rabbit polyclonal anti-mCherry, rabbit monoclonal anti-LINE-1 ORF1p, followed by polyclonal goat anti-rabbit-HRP secondary antibody incubation, hematoxylin counterstaining, mounting, and scanning using a Panoramic Digital Slide Scanner (3D Histech).

3.4. Creation of Plasmid Constructions: The empty *pbiLiv-miR* vector was cloned in a pUC57 plasmid backbone (GeneScript). This includes the bidirectional promoter of the alpha (HADHA) and beta (HADHB) subunits of the human hydroxyacyl-CoA dehydrogenase trifunctional multienzyme complex. The HADHA side of the bidirectional promoter drives the expression of the mCherry fluorescent marker gene, which is disrupted by the first intron of the human eukaryotic translation elongation factor 1 alpha 1 (*EEF1A1*). The mCherry gene is connected to the mouse fumaryl-aceto-acetate dehydrogenase (*Fah*) CDS by a T2A self-cleaving peptide to provide bicistronic expression, then a bGH polyadenylation signal terminates the expression unit. The HADHB side of the bidirectional promoter is followed by a multiple cloning site (MCS) and a bovine growth hormone (bGH) polyadenylation signal. The entire structure is flanked by inverted terminal repeats of SB transposon.

The *pbiLiv-miR-Luc* construct was created by amplifying the firefly Luciferase gene from the pGL3-Basic plasmid (Promega) by PCR and inserting it into the NheI/PacI sites of the multiple cloning site (MCS) in *pbiLiv-miR*. To generate *pbiLiv-miR-mORFeus* variants, the following expression units were inserted into the MCS in the B side of the HADHA/B promoter: TF monomers, ORF1 coding sequence, 3'UTR, hsvTK polyA, EGFP (in reverse orientation), interrupted by the hGamma Globin intron, CMV promoter, and SV40 polyA.
To gain higher cargo capacity, we replaced the SB ITRs with the ITRs of the PiggyBac transposon system. During gene delivery, we co-injected them with the hyperactive piggyBac transposase (hyPBase**).**

3.5. Hydrodynamic Tail Vein Injection: Plasmids were prepared using the NucleoBond Xtra Maxi Plus EF Kit. Mice aged 6–8 weeks received injections of plasmid DNA (4 μg of the transposase helper plasmid (SB100 or hyPBase) and 50 μg of the transposon-containing

plasmid) diluted in sterile Ringer's solution. We administered a volume equal to 10% of body weight through the lateral tail vein in 5-8 seconds to reach a hydrodynamic effect.

3.6. Administration of Screened Compounds for LINE-1 Inductor Properties: Following hydrodynamic injections and NTBC withdrawal, the tested compounds were administered i.p., 2 times a week, for 3 months, with vehicle controls receiving DMSO/corn oil.

3.7. Stereomicroscope Imaging: Whole mouse liver images were captured using an Olympus SZX12 fluorescence stereozoom microscope. The microscope was equipped with a 100-W mercury lamp and specialized filter sets to selectively excite and emit EGFP and mCherry fluorescence.

3.8. *In Vivo* Bioluminescence Imaging: Bioluminescence imaging was conducted with an IVIS Lumina III instrument. Mice were injected with 150 mg/kg luciferase substrate, intraperitoneally and anesthetized using the RAS-4 Rodent Anesthesia System (PerkinElmer) Imaging was started 10 minutes post-injection, capturing emitted photons with exposure times ranging from 1 to 10 seconds. The Living Image software (PerkinElmer) was used to quantify the average radiance within a defined circular region of interest, measured in photons per second per cm^2 per steradian ($p/s/cm^2/sr$).

3.9. Liver Perfusion and Hepatocyte Isolation: Mice were anesthetized and perfused through the *superior vena cava;* hepatocytes were isolated using a multistep collagenase perfusion method. Cells were filtered, centrifuged, and counted while cell viability was evaluated through the trypan blue exclusion test.

3.10. FACS-Based Measurement of EGFP-Positive Hepatocytes: Isolated hepatocytes were analyzed for EGFP fluorescence using a BD FACSAria™ Fusion Flow Cytometer.

3.11. qPCR for Detecting Intronless ORFeus Copies from Retrotransposition Events: qPCR analysis targeted exon-exon junctions of spliced, intronless EGFP. Reactions were conducted in triplicates, and data were analyzed using Rotor-Gene Q software.

3.12. Data Visualization and Statistics: Data visualization was performed using R and GraphPad Prism. Statistical analyses included Fisher's exact test, Pearson's chi-squared test, and independent samples t-tests, with statistical significance established at a threshold of $P < 0.05$.

4 RESULTS

4.1 Investigation of LINE-1 ORF1 Protein Expression in FFPE Human Specimens

LINE-1 ORF1p has gained attention as a biomarker in various cancers, showing expression in malignancies such as breast, ovarian, bladder, esophageal, and pancreatic cancers. Studies suggest that LINE-1 upregulation occurs early in tumor development, as seen in Barrett's esophagus. To explore the phenomenon further, we screened hundreds of tumor samples using TMA technology, observing that ORF1p expression increases with tumor progression.

We were the first to examine ORF1p expression in cervical intraepithelial neoplasia (CIN) and cervical cancer through a comprehensive analysis of 143 cervical specimens. The extent of ORF1p positivity has been associated with the CIN grade. To quantify this, we developed a scoring system based on a modified version of the Immunoreactivity Scale (IRS) by *Remmele and Stegner*. The final IRS score was obtained by multiplying the staining intensity and the extent of epithelial involvement.

ORF1p immunopositivity was detected in the vast majority [110/112 (98.2%)] of dysplastic and neoplastic (CIN and invasive cancer) specimens, whereas 19/24 (79.2%) of normal cervical specimens completely lacked ORF1p expression. The observed pattern of ORF1p expression showed a progressively increasing extent and intensity with advancing CIN grades. CIN I exhibited mild ORF1p expression in the lower one or two-thirds of the cervical epithelium [14/16 (87.5%)], whereas CIN II demonstrated moderate to strong ORF1p expression spanning the lower two-thirds [29/46 (63.0%)]. Pronounced transepithelial ORF1p immunopositivity characterized CIN III cases [13/14 (92.8%)] and cervical cancer [30/32 (93.8%)].

4.2 Measurement of Somatic LINE-1 Retrotransposition Activity in Mice

Measuring LINE-1 retrotransposition in transgenic mice poses challenges due to the early activation of the ORFeus reporter system in germline and embryonic tissues. This early activation causes widespread expression before full somatic development, making traditional transgenic mouse models unsuitable for this purpose.

To address this, we developed a reliable somatic transgenesis platform using a preclinical mouse model of type I tyrosinemia. This model allows stable somatic gene expression

without germline transmission. We knocked out the *Fah* gene in mice by deleting exons 2 to 5 (C57BL/6N-Fah$^{tm1(NCOM)Mfgc/Biat}$), which, without treatment, leads to liver failure. To manage this, we administered Nitisinone (NTBC) to prevent acute toxicity until the mice received gene therapy.

For gene delivery, we used hydrodynamic tail vein injection of the therapeutic plasmid construct with SB100 transposase for stable integration into the hepatocyte genomes. To overcome gene silencing, we linked *Fah* and our transgene of interest using the bidirectional HADHA/B promoter, ensuring a positive selection of FAH-expressing cells. We aimed to achieve permanent transgene expression by replacing $Fah^{-/-}$ hepatocytes with $Fah^{+/+}$ cells by multinodular repopulation, benefiting from the growth advantage of the corrected, FAH-expressing cells This system enabled us to stably express and track transgenes like firefly luciferase (Luc) in *Fah*-deficient mice. Bioluminescence imaging confirmed sustained Luc expression in *Fah* KO mice, while expression declined in wild-type controls. By 3 months, the average bioluminescence intensity in WT mice dropped to an almost undetectable level (387 times drop), while in *Fah* KO mice it reached a 61 times increment compared to the initial level.

Using the insights from these experiments, we could update our genomic technology "toolbox" to be capable of assessing somatic LINE-1 retrotransposition activity in transgenic liver cells of mice. We modified the vector system described above by inserting the ORFeus reporter element into the B side of the HADHA/B bidirectional promoter. We named it *pbiLiv-miR-mORFeus*, which was utilized to generate transgenic livers that permanently express LINE-1 reporter elements. To facilitate genomic integration of the expression cassette, we switched to PiggyBac transposon ITRs and hyperactive transposase helper (hyPBase) to gain high cargo capacity for the transposition of this large-size (~10 kilobases long) expression unit.

We co-injected the *pbiLiv-miR-mORFeus* plasmid with the hyPBase helper plasmid to $Fah^{-/-}$ mice, then we discontinued NTBC administration. In several hepatocytes, successful chromosomal integration was catalyzed by the transposase. Due to the selective growth advantage of these FAH-expressing cells, over 3 months, almost the entire population of hepatocytes expressed the ORFeus reporter. The presence of the FAH and ORF1p expression in the treated $Fah^{-/-}$ livers 3 months post-gene delivery were confirmed by IHC.

4.3 Assessment of Chemical-Induced LINE-1 Retrotransposition Activity

To validate our somatic LINE-1 activity measurement system, we measured the impact of exogenous factors like therapeutic drugs or foodborne chemicals on LINE-1 activity by administering various substances to our reporter-bearing transgenic mice.

After 3 months of drug administration, upon stereomicroscopic macrovisualization, we calculated the number of EGFP-positive hepatocyte colonies in drug-treated mice compared to the non-treated controls. Fluorescence stereomicroscope images highlighted a marked increase in the density of EGFP-positive cell colonies, in the livers of the FICZ (6-Formylindolo[3,2-b]carbazole) treated animals, suggesting that FICZ induces LINE-1 retrotransposition activity in somatic cells. Notably, forced expression of the ORFeus reporter induced retrotransposition events in the untreated control group as well, evidenced by the presence of a small number of EGFP-positive hepatocytes.

To further quantify EGFP-positive cells undergoing retrotransposition, we performed multistep collagenase perfusion to isolate hepatocytes. FACS analysis revealed significantly higher proportions of EGFP-positive cells in Decitabine-treated mice compared to untreated controls, confirming increased retrotransposition activity. Additionally, qPCR analysis targeting intronless EGFP sequences further supported these findings, showing elevated levels of retrotransposition in the treated group. These methods together validated that Decitabine enhances LINE-1 retrotransposition.

5 Discussion

LINE-1 elements are autonomous retrotransposons whose activity is linked to genomic instability. The investigation of LINE-1 activity has gained significant momentum, revealing its role not just as a biomarker but also as a contributor to tumorigenesis. This thesis summarized different important aspects of LINE-1 biology. First, we demonstrated an illustrative example of how a LINE-1-related protein, ORF1p is expressed during carcinogenesis from very early stages. Subsequently, we have developed a novel somatic transgenesis platform by which we could create a viable animal model to assess somatic LINE-1 activity, offering a novel tool for toxicology and cancer research.

Derepressed LINE-1 activity is present in over 50% of tumors and is recognized as a hallmark of cancer, contributing to genomic instability and genetic heterogeneity. Unique

somatic LINE-1 integration events were identified as causative factors of many cancer driver events in multiple types of tumors.

The utility of ORF1p as a pan-tumor biomarker is emerging, not only as an IHC-based marker, but preliminary research has shown that ORF1p is detectable from serum and other body fluids, which suggests its potential in non-invasive cancer screening, early detection, or even treatment monitoring. On the other hand, targeting LINE-1 is possible since ORF1p is absent in normal tissues but abundant in dysplastic lesions. Suppressing LINE-1 activity with 3TC (Lamivudine), an anti-HIV nucleoside reverse transcriptase inhibitor (NRTI), has shown therapeutic benefit in the treatment of metastatic colorectal cancer in a phase 2 clinical study. Drugs that target LINE-1 activity may improve future cancer treatments.

It is now a well-established fact that novel somatic LINE-1 insertions can generate tumor driver events potentially responsible for sporadic cancer cases where no clear cause or known trigger can be identified. We hypothesize that numerous environmental factors can affect LINE-1 activity, although their impact has not yet been investigated. Developing more sensitive methods to detect LINE-1 activity, as explored in this work, could improve our understanding of genotoxicity and help identify risks posed by environmental exposures that might induce LINE-1 activity.

Using *Fah*-deficient livers, we achieved long-lasting and physiological gene expression throughout almost the entire liver of these *Fah*-deficient mice. Our versatile gene expression modification system technology has updated the existing methods by using the hyperactive SB100 transposase for more effective chromosomal transgene delivery. We aimed for tag-free protein expression linked to the selection marker, using the HADHA/B bidirectional promoter for balanced expression. This promoter allows marker-linked expression of untagged native or mutant proteins, achieving a nearly 1/1 transcript ratio, which is crucial for producing protein complexes.

This platform enabled us to insert the synthetic LINE-1 reporter element, ORFeus, to the first side of the promoter, while keeping *Fah* as a positive selection marker gene on the opposite side of the promoter, which facilitates the selection of successfully transfected cells. This configuration ensures that the expression of the LINE-1 reporter is tightly linked to the integration and activity of the transgene within the hepatocytes, enabling the monitoring of LINE-1 activity in a physiologically relevant context.

The implementation of this vector into our *Fah* KO mice allows for the assessment of LINE-1 retrotransposition events. The retrotransposition reporter protein (EGFP) encoded by the ORFeus element is only expressed upon successful retrotransposition, providing a direct readout of LINE-1 activity. The nearly 100 million hepatocytes available in a single experimental animal provide a substantial pool of cells for functional testing. This high number of testable cells allows for the assessment of chemical compounds and their impact on LINE-1 activity as a potential genotoxic effect.

We anticipate our mouse model as a valuable tool in toxicology for evaluating the genotoxic potential of indirect mutagens via the activation of LINE-1 retrotransposons, thus identifying new environmental risk factors and enhancing public health safety.

It is important to acknowledge that the study relies on a mouse model, which presents inherent limitations due to species-specific differences. For example, metabolic rates, immune system responses, hormone regulation, and drug metabolism could differ between human and laboratory mice. The controlled environment in which laboratory mice are kept, does not replicate the diverse and complex environments of humans. Factors such as diet, stress, and exposure to pathogens can influence disease development and progression differently, requiring careful consideration when extrapolating findings. On the other hand, the fundamental biology of LINE-1 elements is largely conserved between mice and humans. Both species exhibit minimal LINE-1 activity in the germline, with somatic activity typically suppressed under physiological conditions. However, during tumorigenesis, LINE-1 activity is reactivated in somatic cells, and the shared mechanisms of LINE-1 control in both species, along with the functional interchangeability of human and mouse LINE-1 proteins, make our LINE-1 activity reporting murine toxicology assay as relevant as conventional rodent toxicology studies.

Overall, the study demonstrates the effectiveness of the ORFeus reporter system implemented in a somatically transgenic mouse model for the detection of somatic *de novo* LINE-1 retrotransposition events and highlights the potential genotoxic effects of chemicals. These findings pave the way for further research into the mechanisms by which environmental factors influence genomic stability and cancer susceptibility.

6 CONCLUSION

Our results highlight the potentially widespread and important role of somatic LINE-1 activity in the development of sporadic cancers. We presented quantitative evidence demonstrating a progressive increase in LINE-1 ORF1 protein expression as cervical neoplasia advances through higher histological grades. The developed *in vivo* technology platform represents a novel approach to evaluate the genotoxic effects of chemicals that might not be detectable through traditional methods. The ability to measure LINE-1 retrotransposition activity in somatic cells provides a new dimension to understand how environmental factors contribute to cancer development. This innovative tool could be used to screen various chemicals for their potential to induce genomic instability, thereby aiding in the identification of new environmental risks and contributing to public health safety.

Key Novel Findings of the Ph.D. Thesis:

1, This thesis demonstrates the potential of LINE1 ORF1p as a biomarker for detecting Cervical Intraepithelial Neoplasia. ORF1p expression was found to be associated with the severity of dysplasia, showing a progressively increasing extent and intensity across the CIN grades.

2, A clinical advantage of using ORF1p is its ability to reliably differentiate between normal cervical epithelium and CIN I lesions. This capability makes ORF1p a valuable marker for early-stage diagnosis and monitoring the progression of cervical carcinogenesis.

3, We developed the first somatically transgenic mouse model for *in vivo* measurement of LINE-1 retrotransposition activity through the creation of the *pbiLiv-miR-mORFeus* plasmid which was hydrodynamically co-administered with the hyPBase transposase to *Fah* knockout mice. This approach successfully overcomes the limitations of previous models, which were hindered by germline activation of LINE-1.

4, The mouse model was utilized to investigate the effects of environmental chemicals on LINE-1 activity. Compounds such as Decitabine and FICZ were shown to increase LINE-1 retrotransposition activity in somatic cells. These findings suggest that these chemicals could pose genotoxic risks through LINE-1 activation.

7 FUNDING

The projects referred to in this thesis were, in part, funded by the Hungarian Academy of Sciences Momentum II Programme (LP2015-5/2015), the GINOP (GINOP-2.3.2-15-2016-00024) and 2021-1.1.4- GYORSÍTÓSÁV-2022-00018 project of the National Research, Development and Innovation Fund, Hungary. R.K. was supported by the ÚNKP-23-3 New National Excellence Program of the Ministry for Culture and Innovation from the source of the National Research, Development, and Innovation Fund, Hungary.

8 LIST OF ABBREVIATIONS

CIN Cervical Intraepithelial Neoplasia
DNA Deoxyribonucleic Acid
EGFP Enhanced Green Fluorescent Protein
FACS Fluorescence-Activated Cell Sorting
FAH Fumaryl-acetoacetate Hydrolase
FICZ 6-Formylindolo[3,2-b]carbazole
FFPE Formalin-Fixed Paraffin-Embedded
HADHA Hydroxyacyl-CoA Dehydrogenase Trifunctional Multienzyme Complex Alpha Subunit
HADHB Hydroxyacyl-CoA Dehydrogenase Trifunctional Multienzyme Complex Beta Subunit
hyPBase Hyperactive piggyBac transposase
IHC Immunohistochemistry
IRS Immunoreactivity Scale
LINE-1 (L1) Long Interspersed Nuclear Element-1
LTR Long Terminal Repeat
NRTI Nucleoside Reverse Transcriptase Inhibitor
NTBC Nitisinone
ORF Open Reading Frame
qPCR Quantitative Polymerase Chain Reaction
RNA Ribonucleic Acid
SB Sleeping Beauty (Transposon System)
SB100 Hyperactive Sleeping Beauty Transposase
TE Transposable Element
TMA Tissue Microarray
TPRT Target-Primed Reverse Transcription
TT1 Type I Tyrosinemia
3TC Lamivudine
WT Wild Type

9 ACKNOWLEDGMENTS

I would like to express my heartfelt gratitude to the many individuals who have contributed to my journey. First and foremost, I would like to thank my supervisor, **Lajos Mátés**, and his partner in both life and profession, **Andrea Mátés-Nagy**, whose guidance and expertise have shaped my research and academic growth. Thank you for welcoming me into your lab back in 2014 as a first-year medical student searching for a mentor and navigating my path toward a research career. During this process, I was lucky enough to have beautiful lab members by my side. Thank you, **Liza Hudoba**, for taking me under your wings and teaching me the ins and outs of research, all while becoming a cherished friend along the way. **András Blastyák,** your endless supply of ideas, creativity, and unrelenting curiosity, combined with your love for science and benchwork, have kept me inspired and grounded in this field. A special thanks goes to **Imre Gergely**; you were always there in stressful times and offered a listening ear whenever I needed it. Our shared sense of humor (*around 4 PM*) brightened even the toughest days. I must acknowledge the contributions of **Ildikó Fekete**, whose technical expertise in the lab was crucial for the success of our experiments. You turned the lab into a true "nest", a place that was always comforting to return to. I want to express my gratitude to **Anna Georgina Kopasz, Dávid Pusztai, Andrea Bakné Drubi, Erzsébet Fehérné Juhász,** and **Khaldoon Sadiq** for their support throughout the years working together. I would like to thank the new members of the lab, **May Raya, Kitti Törőcsik, Katinka Szalmási, Olívia Verebi** and **Dominik Kocsis** for carrying our work forward. Your dedication and enthusiasm ensure that the impact of our research will continue for years to come. I am also grateful to the staff of the **Central Animal House** for their conscientious care and maintenance, ensuring an appropriate environment for our animal experiments. I would like to thank the colleagues from the **Department of Pathology**, Szeged: **Pankotai-Bodó Gabriella, Ivett Csikós, Sarolta Bankó, Ádám Ürmös, Tibor Kun, Beáta Balogh, Krisztián Daru,** and **Mihály Dezső**. My thesis would not have been possible without your contribution. I am outstandingly grateful to **Farkas Sükösd,** who taught me, "*It's not the destination that matters, but the people you share the journey with"*. I wish to express my gratitude to **Ágnes Czibula, Dávid Szüts,** and **Zoltán Villányi** for providing constructive feedback on my thesis. I would like to thank the entire **Biological Research Center**, especially the **Institute of Genetics**, for providing a space where I had the privilege to study molecular biology, surrounded by inspiring minds. Last but certainly not least, I would like to thank my ever-loving **family** and my wonderful **friends**. Your support has meant the world to me.